Language Variation and Change, 3 (1991), 223–239. Printed in the U.S.A.

Contextual conditioning in variable lexical phonology

GREGORY R. GUY
York University

ABSTRACT

The variable lexical phonology proposed in Guy (1991) predicts an exponential relationship among rates of retention in word classes of different derivational histories. A class of words that satisfies the structural description of a variable rule at an early lexical level of derivation will undergo multiple operations of the rule and, therefore, exponentially reduced rates of retention, compared to a class of forms that only satisfies the structural description of the rule at the end of the derivation and thus is subject to its operation only once. For the case of English *-t,d* deletion, it is postulated that monomorphemic words (e.g., *mist*) are exposed to the deletion rule three times in a derivation, whereas semiweak past tense forms (e.g., *left*) are exposed twice, and regular past tense forms (e.g., *missed*) undergo the rule but once.

The present article explores the consequences of this model for other variable constraints on a rule, such as the preceding and following segment constraints on *-t,d* deletion. Word-internal constraints, because they are present throughout the derivation, are shown to have quantitatively different patterns than external constraints, as the latter affect the rule only in its final, post-lexical operation. Four specific quantitative predictions are derived from the model to elucidate this distinction between internal and external constraints, and empirical data are presented to confirm the predictions.

In Guy (1990, 1991) I proposed an approach to phonological variation that combines the variable rule framework with the formal architecture of lexical phonology (Kiparsky, 1982; Mohanan, 1986). The resulting model makes it possible to give precise quantitative predictions about the operation of variable phonological rules on words of differing morphological structures. Briefly, the model predicts an exponential relationship among rates of application for a variable rule operating on word classes that become subject to the rule at different derivational levels in the lexicon and post-lexical phonology. This basic exponential relationship is empirically confirmed in a study of English *-t,d* deletion, which has long been known to apply at different rates to words of different morphological structure (e.g., *mist* undergoes deletion more than *missed*).

This article was prepared with the support of a fellowship from the Stanford Humanities Center, an incomparable venue for the furtherance of intellectual work. The assistance of the Center, and of all my fellow fellows, is gratefully acknowledged.

One consequence of the exponential model not explored in Guy (1991) is the patterning of nonmorphological variable conditioning. Thus, *-t, d* deletion is known to be affected by, among other things, speech style (more deletion in more casual styles), syllabic stress (more deletion in unstressed syllables), and the nature of preceding and following segments (more deletion adjacent to other obstruent segments) (Guy, 1980:5–9). Such constraints are obviously not due to the morphological structure or derivational history of a word, but exemplify the classic kind of independent constraints on a variable process that the variable rule framework is expressly designed to deal with. It is therefore appropriate to test whether such constraints are adequately handled in the variable lexical phonology I propose. In the present article, I show that a family of quantitative predictions about these nonmorphological constraints can be derived from the exponential model of Guy (1991), and that these predictions are empirically confirmed in data from language use.

THE EXPONENTIAL MODEL

In this section, I present a brief sketch of the basic exponential model; for a fuller discussion, readers are referred to Guy (1991). The basic premise of the model is that variable rules applying in a multilevel lexical phonology of the type proposed by Kiparsky (1982) could potentially have multiple opportunities to apply to the forms in a corpus that meet their structural description.

Lexical phonology postulates a single pool of phonological rules operating in a structured derivational process, where they are interleaved with morphological operations. There are minimally three derivational levels or strata: one post-lexical level and two (more, in some versions of the theory) lexical levels, where morphological processes such as concatenation of affixes and compounding apply. In the simplest version of this approach, most derivational affixes and certain irregular inflectional affixes are attached at level one, whereas regular inflectional affixes are attached at level two. At each level, a certain subset of morphological and phonological processes alternately operate. A given phonological rule may operate at more than one level, provided that its structural description is met and there is no specific stipulation preventing it.

The architecture of morphological derivation in this model can be illustrated with the following English examples. Monomorphemic forms (hereinafter abbreviated as M), such as *mist*, *old*, are underived and therefore have their complete phonological shape throughout the derivational process. Semi-weak past tense forms (abbreviated as S), such as *left*, *told*, would undergo affixation of the final stop, irregular internal vowel change, regressive voicing assimilation, and so on, at level one, whereas regular past tense verbs (abbreviated as P), such as *walked*, *missed*, would acquire their affix at level two of the lexicon. Thus, each of these three morphological classes acquires its final segment at a different derivational level.

When variable rules (rules whose probability of application is less than 1.0) are added to this basic architecture, some interesting consequences arise. In a population of forms that fulfilled the structural description of some variable rule early in the derivation (e.g., monomorphemic words subject to *-t,d* deletion), the rule would apply to a certain percentage of them, but fail to affect the remainder. These forms would still meet the structural description of the rule and be subject to its operation again at a later level of the derivation, at which point an additional fraction of them will undergo the rule. Thus, with repeated exposure to the rule, the cumulative percentage of such forms that have undergone the rule will increase. By contrast, a population of forms that only come to satisfy the input conditions for the rule at the end of the derivation (through the operation of some late morpholexical process) will be exposed to the rule only once and will show a lower cumulative percentage of rule application. Thus, in the *-t,d* case, regular past tense forms should only have one (post-lexical) exposure to *-t,d* deletion, because their final stops are not attached until the last lexical level of derivation, whereas monomorphemic words would be subject to the rule three times in a three-level model – once at each level.

If we further assume that variable rules have a fixed basic rate of application (an "input" probability; Cedergren & Sankoff, 1974:339) regardless of level of operation, then this model predicts that the rates of retention in derivationally differentiated morphological categories should show an exponential relationship. That is, if x is the fraction of forms that do *not* undergo the rule in a single application, then only x^2 will remain unaffected after two applications, and x^3 after three applications. For example, a rule with a probability of application of .5 should show retention of .50, .25, and .125 of the forms in morphological categories subjected to one, two, and three applications, respectively. In the *-t,d* case, the model predicts that retention in monomorphemic words should be approximately the cube of the retention rate in regular past tense forms, and the retention rate in semiweak verbs (which undergo affix attachment at level one) should be about the square of the regular past tense rate.

These quantitative predictions, which I call the exponential model, are empirically confirmed in the results reported for seven speakers in Guy (1990, 1991) and in various other data sets reported by other researchers. Table 1 illustrates the exponential relationship with excerpts from the findings of Guy (1991) and from Santa Ana (1991), which is a study of 45 Chicano English speakers. Other corpora that confirm the model may be found in Roberts (1991), a study of 17 children, and Wolfram's (1969) study of 60 speakers, which I examine in more detail later.

In Table 1, the middle column, labeled "estimated p_r", gives the value for the base rate of retention that is derived from observed retention in each morphological category. This is calculated by taking the cube root of the reten-

TABLE 1. *The exponential relationship:* -t,d *retention in two data sets*

	N	% Ret.	Estimated p_r	Significance
Guy (1991) (7 speakers)				
Monomorphemic	658	61.9	.852	Best-fit $p_r = .85$
Semiweak past	56	66.1	.813	Chi-square = 1.28, $p = .55$
Regular past	181	84.0	.840	
Santa Ana (1991) (45 speakers)				
Monomorphemic	3724	42.1	.7494	Best-fit $p_r = .75$
Semiweak past	297	59.3	.7698	Chi-square = 1.17, $p = .57$
Regular past	836	74.3	.7428	

tion rate in monomorphemic words and the square root in semiweak past tense forms; in regular past tense forms, it is simply equal to the observed rate of retention. It can be noticed that these three independent estimates of the base rate are very close to one another. From these, I have calculated a best-fit estimate of this value, and expected values of retention and deletion. The results show convergence between the data and the predictions of the exponential model. A chi-square test shows the model is not rejected; on the contrary, it has a very high p value.

NONMORPHOLOGICAL CONSTRAINTS: QUANTITATIVE PREDICTIONS OF THE MODEL

Complications to this basic picture emerge when we consider that variable rules are normally conceived of as being affected by multiple independent variables, as mentioned earlier. Thus, the likelihood of a rule applying in a particular case is a function not only of the input probability, but also of other variable conditions. This function is usually expressed, in the current standard version, by the logistic equation in **(1)**:

$$\frac{P_{ijk}}{1 - P_{ijk}} = \frac{P_0}{1 - P_0} \times \frac{P_i}{1 - P_i} \times \frac{P_j}{1 - P_j} \times \frac{P_k}{1 - P_k} \qquad \textbf{(1)}$$

In this formulation, P_{ijk} is the probability of the rule applying to a form that is in the context of conditioning factors i, j, and k; P_0 is the input probability; and P_i, P_j, P_k are the factor weights or conditional probabilities associated with the conditioning factors i, j, k.

The factors that may appear at a particular point in this equation are assembled into factor groups, representing a particular point of conditioning in the rule with all of the values that can occur at that point. Thus, the factor i in the equation given might represent the effect of a following segment on *-t,d* deletion, with possible realizations of consonant (which we might represent as i_1), vowel (say i_2), and pause (i_3). Collectively, they form a factor group.

Previous variationist treatments of *-t,d* deletion (such as Guy, 1980) treated morphological category as one of the conditioning factors in such a multivariate model. This is tantamount to claiming that rates of application in the three morphological categories of M, S, and P are independent of one another (as factor values in a group are not mutually constrained, except that they are collectively adjusted so as to average to .5). But Guy (1991) showed that these rates are more efficiently accounted for by the exponential model, which fits the data as well as the logistic model but uses fewer parameters. Each speaker's values for the three categories are predicted from a single estimate of the base rate of rule application (i.e, p_0), whereas the logistic model requires an additional parameter for each category.

The exponential model therefore provides a structural explanation for the quantitative patterning of the morphological conditioning on *-t,d* deletion and removes such conditioning from the list of factor groups whose effect on the rule should be represented by the logistic model. However, the exponential model does not alter in any way the basic rationale or analytical technique of the variable rule framework insofar as nonmorphological conditioning is concerned. These, I assume, continue to operate according to the standard logistic model. Nevertheless, if the exponential model is correct, there are certain quantitative consequences for such nonmorphological conditioning factors. These issues of context-sensitivity were ignored in Guy (1991). The remainder of this section develops the quantitative predictions that the model makes for such conditioning factors.

Quantitative predictions of the exponential model: Internal constraints. If the *-t,d* deletion rule is applying more than once to certain classes of words, but preserving the same rule form, constraints, and factor weights at each application, there should be several quantitative consequences discernible in the data. Most importantly, word-internal constraints on the rule (such as the nature of the preceding segment, the stress level of the final syllable in the word, etc.) should be differentiated from external, cross-word boundary effects (such as the effect of a following segment), because the former are available to condition the rule at every operation, whereas the latter are present only at the single post-lexical operation. Additionally, within each set of constraints, there will be certain perturbations of the basic exponential relationship among the derivational classes. In this section, I sketch the quantitative implications for the internal constraints, beginning with an informal overview, followed by a more precise mathematical treatment. The next section gives a similar account of the predictions made by the model for external constraints.

Informally speaking, when we examine a population of tokens, the effect of internal constraints on a variable rule such as *-t,d* deletion should be apparently magnified for those words subject to repeated rule applications. An internal constraint that is conservative in nature (i.e., one that disfavors application, such as the presence of a preceding liquid) will exercise its conservative restraint on each of three applications of the rule to a population of monomorphemic words, but only once on a set of past tense verbs. Similarly, a constraint that promotes rule application, such as the presence of a preceding stop, will promote it repeatedly in the one morphological class but not in the other. Therefore, the *range* of values in a factor group (the spread between highest and lowest value) should be greater for words subject to repeated applications than for words subject to only a single application.

Despite this amplification of the internal constraints in repeat applications, the basic exponential relationship (among retention rates in words exposed to different numbers of rule applications) should be clearly preserved within tokens with the same internal factor. Thus, if we look only at words with a preceding stop, we should find that the rate of retention in monomorphemic words is the cube of the retention rate in regular past tense verbs, and that the rate in semiweak verbs is the square of this value. The same should occur when examining only words with a preceding liquid or just stressed words or any other set showing a constant internal factor.

The mathematical reason for these two predictions can be discerned from the logistic equation given in (1). If we hold constant a particular value of some internal constraint p_i, in addition to the constancy of the value of p_0 that the exponential model postulates, it will be seen that the product of the two terms $\left(\frac{p_0}{1-p_0}\right)$ and $\left(\frac{p_i}{1-p_i}\right)$ will be constant at each application of the rule. Therefore, the exponential relationship among words with the same internal factor, but belonging to different morphological categories, is unaltered by the presence of the internal factor weight in the equation. Hence, we can formulate Prediction I.

Prediction I. For internal constraints, words showing the same constraint, but differing in derivational history, will preserve the exponential relationship (i.e., $M = P^3$, $S = P^2$; or conversely, $\sqrt[3]{M} = \sqrt{S} = P$).

By the same token, however, this preservation of the exponential relation within a sample of words showing the same internal factor implies that the difference between them and another sample with a contrasting internal factor must increase with repeat applications. It should be clear that the constant product of the two terms $\left(\frac{p_0}{1-p_0}\right)$ and $\left(\frac{p_i}{1-p_i}\right)$ can be translated into a constant predicted value of retention for each application, which we can designate as R_i. If we compare this with another predicted value in the context of a different internal factor (say, R_k), then with multiple applications, the proportional difference in their surface rates of retention must increase. For example, if R_i is .8 and R_k is .6, then for words exposed to the rule only once, the ratio between their surface rates of retention is .8/.6 = 133%; in other words, the more conservative environment shows 33% more retention than the less conservative one. But after two applications the ratio increases to .64/.36 = 178%, and after three applications to 241%. The proportional difference between factors in an internal factor group must therefore appear larger in words exposed to more operations of the rule. Thus, we make Prediction II.

Prediction II. For internal constraints, the range of surface retention rates in a factor group will increase in derivational categories subject to multiple applications. As a corollary, Varbrul factor-weight estimates for internal constraints that are based on subsamples of single derivational classes differ in the same way: the range of values will be smaller for P words than for S or M words.

Quantitative predictions for external constraints. Constraints on a variable rule that are external to the word must, in a lexical phonology, operate only at the post-lexical level. Therefore, they do not affect the lexical derivation of the word and should not be sensitive to word-internal morphological structure. In the *-t,d* case, this means that the following segment effect, which refers across word boundaries to the first segment of a following word (if any), should be constant for all morphological classes. Each class is subject to these constraint effects but once, in the post-lexical phonology. This observation may be summarized as Prediction III, which contrasts with the internal pattern of Prediction II.

Prediction III. For external constraints, the range of surface retention rates in a factor group will be the same for all derivational categories. As a corollary, Varbrul factor-weight estimates for external constraints that are based on subsamples of single morphological classes should show approximately the same range of values for P, S, and M words.

Finally, the constancy of external effects across the morphological classes implies that the exponential relationship among these classes will be systematically distorted when one examines data drawn from a subsample of tokens that all show the same external factor. For example, if we looked only at words with a following obstruent, we should find that the rate of retention in monomorphemic words is not equal to the cube of the retention rate in regular past tense verbs, and that the rate in semiweak verbs is not equal to the square of this value. The direction of the difference between the observed rates and the predicted exponential values will depend crucially on whether the external constraint in question favors or disfavors the rule. For an external factor that promotes rule application, retention in forms subject to multiple applications will be greater than the powers of the rate of retention in forms subject only to a single application ($M > P^3$, $S > P^2$). For an external factor that hinders rule application, the rate of retention in forms subject to multiple applications will be less than the powers of the rate in single-application forms ($M < P^3$, $S < P^2$).

The reason for this departure from the exponential relationship is precisely that the favoring or disfavoring effect of the external constraint is only present in the last, post-lexical application of the rule and does not iterate in the lexicon. Therefore, the portion of the final surface retention rate that is due to this effect is not exponentially repeated.

A mathematical formulation of this relationship may be given as follows. For a derivational category subject to a rule only once (e.g., regular past tense forms in the *-t,d* case), let P represent the true base rate of retention (which is equal to 1 minus the input probability for the rule). In a pool of such forms that are observed in a conservative (e.g., pre-vocalic) position, let the observed surface rate of retention be represented as P'. Under the conditions outlined here, P' does not provide a good estimate of P, but rather is actually boosted above that level by the effect of the conservative context; that is, $P' = P + x$. If we now examine words subject to the rule twice (i.e., semiweak verbs in this case), the predicted rate of retention should be $P(P + x)$. This is clearly less than the square of P', because $P(P + x) < (P + x)^2$ for all positive values of x. Therefore, if our point of departure for estimation is P', the observed surface rate of retention in this conservative context, then systematically we will find that surface retention rates in the semiweak class (S) should be less than the square of P'. Similarly, surface retention rates in the monomorphemic class (M) will be less than the cube of P'.

For environments that promote deletion (such as preconsonantal position for *-t,d* deletion), all the inequalities in the argument just presented are reversed. P' can now be symbolized as equal to some reduction from the true retention rate (i.e., $P' = P - x$), and then $P(P - x) > (P - x)^2$. In the *-t,d* case, this means that observed retention rates in the monomorphemic and semiweak classes will be greater than the cube and square of the surface rates in the regular past tense verbs. These relationships can be summarized as Prediction IV.

Prediction IV. For external constraints, words showing the same constraint, but differing in derivational history, will deviate from the exponential relationship. When the external constraint favors rule application, retention rates will exceed an exponential relation (i.e., $M > P^3$, $S > P^2$), whereas for disfavoring constraints, retention rates will fall short of the exponential relation ($M < P^3$, $S < P^2$).

EMPIRICAL TESTS OF THE PREDICTIONS

If the exponential model accurately represents what is happening in language production, then Predictions I–IV should be empirically verifiable. As always, such verification is subject to problems of sample size, statistical fluctuation, and so forth, but the usual methods are available to us to combat such problems (statistical tests, enlarging the sample, replicating studies, etc.). In this section, I report data relevant to these predictions drawn from the same corpus on which Guy (1991) is based (see Appendix for details) and from published data on *-t,d* deletion appearing in Wolfram (1969). Wolfram's study is based on the Detroit dialect survey corpus of sociolinguistic interviews (Shuy, Wolfram, & Riley, 1968). Wolfram (1969) examined a socially stratified group of 48 black speakers (and a comparison group of 12 upper middle-class white speakers). In his treatment of *-t,d* deletion, he did not explicitly report *N*s, although from his description of methods (p. 58), it can be inferred that for each of the five class and racial subgroups he described, there should be approximately 240 tokens of monomorphemic words and 180 tokens of "bimorphemic" words. This bimorphemic class may not be perfectly comparable with the regular past tense (P) class in my analysis, because it may include tokens of semiweak past tense forms; Wolfram's discussion is unclear on this point. But even so, it should be heavily weighted toward the regular forms.

Data

Internal constraints. For purposes of this test of Predictions I and II, I consider only the most prominent internal constraint on *-t,d* deletion: the nature of the preceding segment. Detailed analyses of this effect usually give separate treatments of sibilants, nonsibilant fricatives, stops, nasals, and liquids but do not always find significant differences among all these and tend to find some instability in the factor values and ordering of this factor group. To conduct a clear test of the exponential model, robust *N*s and stable values are more important than phonetic detail; therefore, I adopt a less detailed analysis, combining nonsibilant fricatives and stops in an obstruent category. Even so, *N*s for the semiweak verb category become unreliably small: all four preceding environments showed expected values of less than 5 for either retentions or deletions, or both, making the chi-square test unreliable. Accordingly, detailed calculations are not presented in this section for semiweak verbs.

According to Prediction I, if the preceding segment is held constant, the exponential relationship among derivational categories M, S, and P should be preserved. Table 2 shows the test of this hypothesis for the M and P categories based on the data from the Guy (1991) corpus. The table shows *N*s, percentages of retention, and the values of p_r (the base rate of retention) estimated according to the exponential model (i.e., the cube root of the surface retention rate for M words and the surface rate itself for regular past tense forms). In the second half of the table a best-fit estimate[1] of p_r is given, along with a chi-square test of the exponential hypothesis for each constant preceding environment.

TABLE 2. *Preceding segment constraint on* -t,d *deletion: Retention rates and estimates of base rate of retention according to the exponential model (Guy, 1991 corpus)*

	Morphological class					
	M			P		
Preceding segment	Retention/Total	%	Est. p_r	Retention/Total	%	Est. p_r
Sibilants	134/269	49.8	.7927	31/40	77.5	.7750
Obstruents (stops and nonsibilant fricatives)	24/34	70.6	.8904	79/93	84.9	.8495
Nasals	133/214	62.2	.8534	16/19	84.2	.8421
Liquids	116/141	82.3	.9370	26/29	89.7	.8966

		Expected retention			
	Best p_r	(M)	(P)	Total chi-square	$p =$
Sibilants	.790	132.6	31.6	.0835	.87
Obstruents	.865	22.0	80.4	.7050	.75
Nasals	.853	132.8	16.2	.0185	.94
Liquids	.932	114.2	27.0	.7318	.78

The results show that the retention rate in regular past tense verbs closely approximates the cube root of the retention rate in monomorphemic words. The chi-square test shows that the exponential model for these two categories is not rejected at the .05 level for any of the four preceding contexts; in fact, the high p values indicate a good fit between data and model.[2] These data are therefore consistent with Prediction I.

Prediction II is also confirmed by the data. This effect can be seen by examining the retention frequencies in Table 2. For monomorphemic words, the range of retention rates is from a low of 49.8% following a sibilant to a high of 82.3% following a liquid. This gives an absolute range for the factor group of 32.5% and a proportional difference of 1.65. But for the regular past tense verbs, the range of values runs only from a low of 77.5% to a high of 89.7%, an absolute range of 12.2% and a proportional difference of 1.16. This confirms the prediction that internal constraints will be magnified on the surface by repeat applications in the lexicon. However, using these raw frequencies is potentially misleading, as they depend in part on p_0. (They will be minimized for values of p_0 close to 0 or 1 and maximized for p_0 values close to .5.) Therefore, another strategy for testing Prediction II is to compare the Varbrul factor values for the preceding segment factor group that are obtained by analyzing the data from each morphological category separately. These values, given in Table 3, also show the increased range of the effect in the monomorphemic data. The range of the factor

TABLE 3. *Preceding segment constraint on* -t,d *deletion: Varbrul factor weights for separate analyses of morphological classes (Guy, 1991 corpus)*

	Morphological class	
Preceding segment	M	P
Sibilants	.66	.67
Obstruents (stops and nonsibilant fricatives)	.49	.46
Nasals	.59	.41
Liquids	.27	.44
Range:	.39	.26

weights in the preceding-segment factor group is .39 for the monomorphemic words analyzed alone, but only .26 for the past tense forms analyzed alone, further confirming the indications of the raw frequency data.

The predictions of the exponential model can also be tested for the data reported in Wolfram (1969). An adaptation of Wolfram's results for preceding segment are reported in Table 4, where it can be seen that Prediction I is essentially confirmed. The last four columns in the table compare, for each constant preceding environment, the estimate of p_r obtained from the cube root of retention rates in monomorphemic words with the estimate obtained from the base retention rate in bimorphemic words; the pairs of values are systematically close. However, because Wolfram did not report actual *N*s, a statistical test of the model (such as the chi-square test reported for the Guy corpus) cannot be conducted for these data.

Wolfram's data do not confirm Prediction II, however, mainly because he tested only two (for past tense verbs, three) distinct preceding environments and found no significant difference between them. Therefore, no range differences for this factor group are discernible. This result may arise because of the analytical categories Wolfram adopted. He treated all fricatives together, against all sonorants (nasals plus liquids). Most other studies have shown a significant difference in effect between sibilant and nonsibilant fricatives and between nasals and liquids. Thus, the relationship of Prediction II may simply have been obscured by the way these data were partitioned.

External constraints. The main external constraint on *-t,d* deletion that is relevant to Predictions III and IV is the nature of the following segment. Prior studies are unanimous that a following consonant promotes deletion more readily than a following vowel. More detailed studies (e.g., Guy, 1980; Nesbitt, 1984; Labov, 1989; Roberts, 1991) have generally shown a finegrained differentiation of following contexts along a sonority scale, such that

TABLE 4. *Preceding segment effect on* -t,d *deletion: Retention rates and estimates of* p_r *(Wolfram, 1969 corpus)*

Social group	M		P		Estimates of p_r			
					Nasal/Lateral		Spirant	
	Nasal/Lateral	Spirant	Nasal/Lateral	Spirant	M	P	M	P
UMW	87.3	93.1	97.0	96.2	.956	.970	.976	.962
UMN	80.3	88.0	96.4	88.1	.929	.964	.958	.881
LMN	57.1	63.2	85.3	79.4	.830	.853	.858	.794
UWN	28.4	36.1	75.0	68.5	.657	.750	.712	.685
LWN	31.4	24.5	59.3	65.7	.680	.593	.626	.657

deletion rates fall in the order: obstruents > liquids > glides > vowels.[3] The place of pause on this hierarchy has been shown by Guy (1980) to be dialectally arbitrary; it is sometimes a favoring context for deletion, sometimes disfavoring. The most common pattern locates the pause effect at an intermediate value between those of obstruents and vowels.

For present purposes, just as in the case of the internal constraints, large *N*s are more important than phonetic detail, so I do not attempt to distinguish liquids and glides from other consonants and confine the discussion to just the M and P morphological categories. Wolfram did not report separate data on following pause, combining it with vowels in a "nonconsonantal" category.

Prediction III states that in an external factor group the range of surface retention rates and Varbrul factor values should be essentially the same for all morphological categories, with the proviso noted earlier that when retention rates approximate 0% or 100%, the surface rate ranges will shrink and Varbrul values will be more accurate. As can be seen in subsequent tables, in both the Guy (1991) and Wolfram (1969) corpora, retention rates above 90% or below 10% occur. Therefore, I initially illustrate this effect with the Varbrul results for a three-factor analysis of the following segment effect in M words and P words in my corpus, given in Table 5. The ranges obtained for the two classes analyzed separately are virtually identical, in contrast to the ranges shown in Table 3 for the internal constraint. (The raw frequency data, which appear in Table 8, show the range compression that arises when some contexts approach categorical application.)

Wolfram's data, shown in Table 6, also display uniformity of range for this external constraint. Of the 10 ranges shown, 7 fall between 42.1% and 56.3%; the three remaining cases with smaller raw frequency ranges all involve one end of the range falling within 6.5% of categoricality. The average range for this factor group across all five social groups is 41.6% for monomorphemes and a nearly identical 42.9% for past tense forms. With-

TABLE 5. *Following segment constraint on* -t,d *deletion: Varbrul factor weights for separate analyses of morphological classes (Guy, 1991 corpus)*

	Morphological class	
Following segment	M	P
Consonants (including liquids and glides)	.73	.65
Vowels	.31	.24
Pause	.45	.63
Range:	.42	.41

TABLE 6. *Following segment effect on* -t,d *deletion: Percentages of retention and ranges (Wolfram, 1969 corpus)*

Social group	M			P		
	Consonantal	Nonconsonantal	Range	Consonantal	Nonconsonantal	Range
UMW	33.6	88.5	54.9	62.8	97.2	33.4
UMN	21.1	77.4	56.3	50.8	93.2	42.4
LMN	13.3	56.7	43.4	38.3	86.7	48.4
UWN	6.5	34.6	28.1	27.5	75.7	48.2
LWN	2.7	27.9	25.2	24.0	66.1	42.1
Mean ranges:			41.6%			42.9%

out the *N*s, a precise Varbrul analysis cannot be conducted on these data, but a modeling analysis of Wolfram's percentage figures with equal cell sizes yields factor weights of .77 and .23 for following consonantal and nonconsonantal contexts respectively in M words, and .75 and .25 in P words. Such results all conform to what would be expected if the following segment constraint only operates once on words of all morphological categories.

Finally, Prediction IV can also be tested against Wolfram's data. The prediction states that words showing the same external constraint, but differing in derivational history, will systematically deviate from the exponential relationship. For the *-t,d* deletion case, in deletion-favoring environments (such as following consonant) an estimate of p_r based on the cube root of deletion in M words should be greater than one based on the deletion rate in P words (because the deletion-promoting influence of the following consonant has been exercised only once in both cases, rather than iterating throughout the derivation). For deletion-inhibiting environments, such as following vocalic contexts, the M-based estimate should be less than the estimate obtained from P words.

TABLE 7. *Following segment effect on* -t,d *deletion: Estimates of* p_r *(Wolfram, 1969 corpus)*

	Estimates of p_r by constant following context:					
	Following consonantal			Following nonconsonantal		
Estimates based on:	$\sqrt[3]{M}$		P	$\sqrt[3]{M}$		P
Social Group						
UMW	.695	>	.628	.960	<	.972
UMN	.595	>	.508	.918	<	.932
LMN	.510	>	.383	.828	<	.867
UWN	.402	>	.275	.702	<	.757
LWN	.300	>	.240	.653	<	.661

TABLE 8. *Following segment constraint on* -t,d *deletion: Retention rates and estimates of base rate of retention according to the exponential model (Guy, 1991 corpus)*

	Morphological class					
	M			P		
Following segment	Retention/Total	%	Estimated p_r	Retention/Total	%	Estimated p_r
Consonantals	132/305	43.3	.7564	63/83	75.9	.7590
Vowels	168/204	82.4	.9373	72/76	94.7	.9474
Pause	107/149	71.8	.8955	17/22	77.3	.7727

These predictions are systematically confirmed in Wolfram's data, as shown in Table 7. For every social group, the M-based estimates are greater than P-based estimates in the favoring environment for deletion, and less in the disfavoring environment. The inequality signs in the table are all oriented in the predicted direction.

Although these results are gratifying to the theorist, the data from the Guy (1991) corpus do not confirm Prediction IV quite so neatly, as can be seen in Table 8. For following vowels and consonantals, the exponential relation seems to be preserved. The estimates of p_r based on M and P are close, and the small deviations that do occur follow the predicted direction in the case of vowels but actually go slightly in the unpredicted direction for consonantals. The following pause data could be said to strongly support the prediction, provided that pause is a deletion-promoting environment. Because its effect is intermediate between the other two categories, it is difficult to confirm this in any absolute sense, and because it has the smallest *N*s, it would be unwise to make too much of it.

The Guy (1991) corpus thus fails to confirm the predicted inequalities of Prediction IV. How may these results be accounted for within the framework of the exponential model? Two steps are indicated for further refinement of the theory. First, the inequalities may be lent some precision. Given some estimate of the strength of the favoring or disfavoring effect exerted by postlexical constraints, such as the estimates provided by the Varbrul factor weights, it should be possible to quantify the magnitude of the predicted inequalities and apply statistical tests to a robust corpus. And second, the entire process would benefit from a multivariate approach. As it is a basic assumption of the variationist approach that these rules are subject to multiple, simultaneous conditioning from several sources, deriving our estimates while controlling a single factor at a time is inherently limited by the possibility of uncontrolled variation going on elsewhere. It is a useful and necessary first step until better analytical tools are constructed, but a multivariate modeling technique should be a goal for this kind of research.

CONCLUSIONS

These results provide further confirmation of the validity of the exponential model. The model can be extended along these lines to accommodate, in principle, any kind of conditioning factor, and it will make precise, falsifiable predictions that will systematically distinguish between internal factors that participate in iterated applications of a rule and external factors that do not. The data presented here broadly confirm the four quantitative predictions presented. The few failures-to-confirm can all be plausibly accounted for by the relative crudity of the present state of the model and the analytical techniques used here. Further refinements are therefore indicated, but even in its present state the exponential model provides for a level of quantitative precision coupled with explanatory value heretofore impossible in the study of variation.

Furthermore, there is now a substantial body of evidence that an exponential relation exists for English *-t,d* deletion, and that conditions on the rule show patterns consistent with such a relation. An unconstrained, nonexponential model of this process, such as the one incorporated in the Varbrul program, allows for the possibility that retention rates in different derivational categories could fall in ratios like 75%, 50%, 25%, or 80%, 60%, 30%. The facts show that they do not. Rather, we find exponential sequences like 75%, 56%, 42%, or 80%, 64%, 51%. Any alternative to the present model would have to account for these facts, and it is difficult to see how that could be done in a theoretically plausible way without using rule iteration.

More generally, however, these results provide serious empirical evidence for the iterative, multileveled architecture of lexical phonology. They also suggest constraints on that theory, beginning with the requirement that we take it seriously as a model of human language use. In terms of formal structure, they require that the number of iterative lexical levels, at least for English *-t,d* deletion, be limited to two and that the rule applies once at each lexical level, before bracket erasure. It would be an interesting research project to see if these requirements are reconcilable with other proposals on these points, and if they are empirically consistent with other variable processes.

Finally, these results illustrate the mutual utility of quantitative investigation and the construction of linguistic theory. It is lamentable that such a point must still be made in linguistic science, but the dominant practice in the field is to treat theory construction and quantitative study as mutually exclusive—or even antagonistic—enterprises. Therefore, it bears repeating that data and theory have a symbiotic relationship, and that, indeed, this is an elementary requirement for empirical science. Primary studies of language use repeatedly turn up patterns that challenge us in our role as explainers, model builders, theoreticians; similarly, the theories that we construct generate predictions that guide our empirical investigations, helping us to find patterns we might never have suspected otherwise. The exponential model combines empirical observations of production with theoretical proposals about the organization of rule systems to yield refinements to both and to allow novel predictions that, on present evidence, are confirmed.

NOTES

1. These values actually represent a combination of the effect of p_r with the effect of the constant preceding environment, as was discussed in the formulation of Prediction I. Thus, it can be noted that they reflect the usual range of values for this factor group, with liquid most conservative and sibilant most deletion-promoting.
2. When the data for semiweak verbs are included, the goodness of the fit declines, but the model is only rejected in the case of preceding liquids, where p is approximately equal to .04, due to predicted deletions in this cell of 3.68, against observed deletions of 8.
3. Guy (1991) postulated that this order is a consequence of variable surface syllabification involving a rightward attachment of the final *-t,d* as onset of the next syllable. This is completely blocked in the case of obstruents, completely free in the case of vowels, and disfavored in some sequences with following glides. This explanation predicts that the liquids should split, with following /l/ disallowing resyllabification (therefore promoting deletion), while following /r/ allows it. This prediction is confirmed in the Guy (1991) corpus.

REFERENCES

Cedergren, Henrietta J., & Sankoff, David. (1974). Variable rules: Performance as a statistical reflection of competence. *Language* 50:333–355.
Guy, Gregory R. (1980). Variation in the group and the individual: The case of final stop deletion. In W. Labov (ed.), *Locating language in time and space*. New York: Academic. 1–36.
Guy, Gregory R. (1989). *MacVarb application and user documentation*. Stanford: Stanford University, Linguistics Department.

Guy, Gregory R. (1990). *Explanation in a variable phonology:* -t,d *deletion*. Paper presented at NWAVE-XIX, University of Pennsylvania.
Guy, Gregory R. (1991). Explanation in a variable phonology: An exponential model of morphological constraints. *Language Variation and Change* 3:1-22.
Kiparsky, Paul. (1982). Lexical morphology and phonology. In I. S. Yang (ed.), *Linguistics in the morning calm*. Seoul: Hanshin. 3-91.
Labov, William. (1989). The child as linguistic historian. *Language Variation and Change* 1:85-98.
Mohanan, K. P. (1986). *The theory of lexical phonology*. Dordrecht: Reidel.
Nesbitt, C. (1984). *The linguistic constraints on a variable process:* /t,d/ *deletion in Sydney speech*. BA Honours thesis, University of Sydney.
Roberts, Julia. (1991). t/d *deletion in preschool children*. Paper presented at Child Language Research Forum, Stanford University.
Santa Ana, Otto. (1991). *Phonetic simplification processes in the English of the barrio*. Doctoral dissertation, University of Pennsylvania.
Shuy, Roger, Wolfram, Walt, & Riley, W. (1968). *Field techniques in an urban language study*. Washington, DC: Center for Applied Linguistics.
Wolfram, Walt. (1969). *A sociolinguistic description of Detroit Negro speech*. Washington, DC: Center for Applied Linguistics.

APPENDIX

The data on which Guy (1991) and portions of the present work are based were obtained from sociolinguistic interviews with seven native speakers of North American English. These were conducted in 1989 by a research group at Stanford University consisting of Renee Blake, Erica Deese-Dobson, Gregório Firmino, Rudi Gaudio, Martha Porras, Jeff Seinfeld, and Mark Van Haren. The consultant group is diversified on most social dimensions. They range in age from 9-55 and include four females and three males, and four white and three nonwhite speakers. Data-base operations and Varbrul analyses were conducted using MacVarb (Guy, 1989). Data extraction and coding followed standard procedures; see Guy (1991) for details.